Springboard DSC Program

Capstone Project 1 Proposal

Credit Card Fraud Detection

Anonymized credit card transactions labeled as fraudulent or genuine

By Harinee Madhusudhan

October 2019

Problem

Detecting credit card fraudulent transactions.

Who is your client and why do they care about this problem? In other words, what will your client do or decide based on your analysis that they wouldn't have done otherwise?

- The clients are credit card companies. It is important that credit card companies can recognize fraudulent credit card transactions so that customers are not charged for items that they did not purchase.

What data are you using? How will you acquire the data?

- Located some normalized data based on prior analysis. The datasets available from https://www.kaggle.com/mlg-ulb/creditcardfraud, contain transactions made by credit cards in September 2013 by European cardholders. This dataset present transactions that occurred in two days, where we have 492 frauds out of a total of 284,807 transactions. The datasets are highly unbalanced, the positive class (frauds) account for 0.172% of all transactions.

- The datasets contain only numerical input variables, which are the result of a PCA transformation applied to the original data. To preserve anonymity, they cannot provide the original features and more background information about the data. Features V1, V2, ... V28 are the principal components obtained through PCA transformation.

Briefly outline how you'll solve this problem. Your approach may change later, but this is a good first step to get you thinking about a method and solution.

- The objective is to identify the appropriate model. The model could be trained using the data available. Once the model is trained, we can use it to predict future model given the context.

    What are your deliverables?

- All Jypyter notebooks that will be developed

- A required Final Report

- A required presentation slide deck